

The Application of Substrate Theory in Research on the Ethnic Languages of China

著者(英)	Qingxia Dai
journal or publication title	Senri Ethnological Studies
volume	75
page range	281-289
year	2009-09-18
URL	http://doi.org/10.15021/00002569

The Application of Substrate Theory in Research on the Ethnic Languages of China

Dai Qingxia

Central University for Nationalities, China

1. Chinese Researchers' Understanding of Substrate Theory
2. The Substrate Theory Applied to Ethnic Languages in China
3. Difficulties and Challenges that the Substrate Theory Encounters
4. How to Strengthen the Establishment of the Substrate Theory

1. Chinese Researchers' Understanding of Substrate Theory

1.1 China is a multi-ethnic and multi-lingual country, with 56 officially-designated ethnic groups using more than 80 languages. The languages belong to five language families: Sino-Tibetan, Altaic, Austronesian, Austro-Asiatic, and Indo-European. Many languages have been in contact with one another for a long time, bringing about mutual influence and language mixing to different degrees and in various forms. Consequently, there are loan elements due to language contact, substrate elements due to language mixing and language shift, and shared elements due to genetic relationship. How should these elements be distinguished? In the past five decades, linguists have made great efforts to suggest different theories and methods to solve the problem. The substrate theory is one theory that Chinese linguists have adopted to account for language mixing.

1.2 The idea of a linguistic substratum was introduced as early as the beginning of the 19th century. For example, J. H. Bredsdorff, the Italian linguist, put forward the concept in 1821 in his masterpiece "On the Causes of Linguistic Change" ("om Aarsagerne Tilspiogenes Fornandringer") (Chen 1969: 1). Subsequently the concept was endorsed by other linguists. In 1955, the Institute of Language, Soviet Academy of Sciences held a special conference in Leningrad to discuss the issue of substrate theory. At the meeting, many experts expressed their views on the theory and engaged in enthusiastic discussions. A collection of papers was published afterwards. Chinese linguists soon learned about the event, and paid much attention to the substrate theory.

1.3 It was in March 1960 that a definition of “substratum” appeared in *Linguistic Terminology*, edited by the Language Teaching and Research Office of Peking University: “Substratum is a linguistic term about language mixing. Some scholars believe that when an ethnic group abandons its mother tongue and adopts another language, they often bring along old linguistic practices into the new language, establishing a layer of linguistic phenomena beneath the new language. These linguistic phenomena, including various aspects of pronunciation, vocabulary and grammar, are different from the original phenomena of the new language. This group of linguistic phenomena transferred to and existing beneath or behind the new language are known as the substratum.” (Language Teaching and Research Office of Peking University 1960: 25) At that time, this entry was an introduction to substrate theory, and the definition of “substratum” as a basic concept in language contact still looks rather scientific even now.

1.4 In particular, we should mention that Gao Mingkai used the term substratum in the “Language Replacement” chapter of his book *On Language* (1995). Gao’s description was brief but rather close to the facts. He pointed out: “Certain pronunciation habits of an ethnic group can be kept when the people switch to another language, constituting a ‘substratum’ in the latter language. Not only pronunciation but also grammar and vocabulary can produce the phenomenon of substratum.” He wrote: “It is an undeniable fact that substratum appears.” “However, up to now linguists have not arrived at a consensus as to what exactly a ‘substratum’ is and what its characteristics are.” (Gao 1995: 507–509) The book also cited some examples of a substrate in languages. Forty-odd years have passed and we can see that Mr. Gao had a sound grasp of the term and his comments were very much to the point.

1.5 In the last two decades, a remarkable paper was Ouyang Jueya’s “Relationships between ethnic languages and Chinese: An application of substrate theory” (1991). In an in-depth discussion of substrate theory based on the actual linguistic situation in China, Ouyang commented: “Generally speaking, the impact of Chinese on ethnic languages is mostly shown in borrowing. Ethnic languages have borrowed a lot of words from Chinese since ancient times, simultaneously causing changes in their pronunciation and grammar. On the other hand, the impact of ethnic languages on Chinese shows up mostly as a substratum. This situation is closely related to the fact that the ethnic languages spoken in a considerable number of locations have been assimilated by Chinese. It is well-known that northern Chinese dialects are very different from the southern ones. For this there are many reasons, among which the impact of ethnic languages is very important.” (Ouyang 1991: 23–29)

1.6 Based on research results about substrata in the past, we can explain the meaning of the term as follows: “Substratum refers to elements of a previous language kept in another language when an ethnic group or a community has stopped using the former and adopts the latter as the native language”. The previous language is the substratum language, and the “new” language, including elements of the previous language (substrate) and the adopted language, is referred to as a language with substratum elements.

Several points of the explanation given above can be highlighted here:

(1) Substratum is the product of language shift. Without language shift there would not be a

substratum.

- (2) Substratum refers to the elements of the previous mother tongue added to the adopted language (or the target language) when language shift takes place. However, not every language spoken as the new mother tongue contains elements of the old one. In fact, most new mother tongues do not have a substratum.
- (3) Different from both languages, i.e. the old mother tongue and the target language adopted as the new mother tongue, the language with substratum elements, comprising elements from two languages, is a new linguistic system.
- (4) Substratum languages make up a topic for study within the fields of both historical linguistics and sociolinguistics. The two terms “language with substratum elements” and “mixed language” are closely related to each other. A language with a substratum can be treated as a mixed language, and a mixed language usually has a substratum or substrata from other languages.

2. The Substrate Theory Applied to Ethnic Languages in China

2.1 Why have Chinese linguists studying ethnic languages accepted the substrate theory and attached great importance to it? Why has it become a notion frequently mentioned to account for China’s linguistic relations? In my opinion, the reason is that the historical development of China’s languages is very complicated, having various forms of language contact and mixing that are difficult to delineate. Existing in a multi-ethnic and multilingual country, many languages in China have been in contact with others for a long time, influencing others and being influenced. As a result, there are loan elements due to language contact, substrate elements due to language mixing and language shift, and shared elements due to genetic relationship. To identify account for these phenomena, concepts and theories should be put forward, and substrate theory can adequately deal with cases of language contact and language shift.

2.2 In the past 50 years or so, Chinese scholars have utilized substrate theory to study minority languages in China, and identified a number of substratum languages. The following lists previous studies in chronological order.

2.2.1 In “Elements of Ancient Yue in Cantonese”, Ouyang Jueya made use of substrate theory to discuss the relation between Cantonese and Old Yue and concluded that there exists a substratum of Old Yue (or Tai) in Cantonese (Ouyang 1980: 605–616). He continued to publish a series of articles afterwards to offer more examples of the substratum.

2.2.2 In “Numbers in the Ainu Language, with Remarks on the Nature of Ainu” (1981), Zhao Xiangru and Aximu stated, “The Ainu language originally belonged to the Iranian branch of the Indo-European family. Having been affected by Uyghur, Ainu people now use a mixed language among themselves. The vocabulary of this mixed language at first originated mostly from the Iranian branch, but has undergone a great change; and its grammatical structure is based on Uyghur.” The language that Ainu people currently use is actually a variety of

Uyghur with Iranian elements as a substratum, although the authors did not specifically use terms of the substrate theory (Zhao 1981: 44–48).

2.2.3 In “On the Affiliation of the Tai Family of Languages” (1983), Luo Meizhen clearly pointed out that there is a Malay substratum in Tai [or Kam-Tai] languages. She said, “As early as the time of Proto-Sino-Tibetan (PST), i.e., before its evolution into different languages, the ancestors of peoples that now speak Tai languages today renounced their original language and adopted Sino-Tibetan because of ST’s powerful influence. After PST split into individual languages, Tai peoples evolved their own languages. That is why today’s Tai languages share many similarities with ST languages in pronunciation, vocabulary and grammar, but few similarities with Malay. Accordingly, the development of Tai languages is based on ST, and they should be treated as a part of ST. As for the Malay remnants or ‘substratum’ left in Tai languages, there is insufficient evidence to trace their genetic relation with other languages.” (Luo 1983: 30–40)

2.2.4 In “Hezhou Chinese and Altaic Languages”, Ma Shujun analyzed the impact of Altaic on the Chinese spoken in Hezhou, suggesting that a substratum is involved in the formation of Hezhou speech. The Hezhou vernacular as spoken by both Han and Hui (or Muslim) people in Linxia, Gansu, is a variety of Chinese having the same or similar characteristics with Altaic. What is the cause for the similarities? Is it language impact or substrate deposit? After comparing the similar elements in question, the authors concluded, “Generally speaking, if a people are in a weaker position in politics, economy, culture and population, but enjoys population superiority in certain local areas, when it comes to adopting another language, more substrate elements from the discontinued language will be kept. It seems that the formation of the Hezhou vernacular has followed this path.” (Ma 1984: 50–55)

2.2.5 In “On the Voicing of Voiceless Fricative Initials in Languages in Southern China” (1990), Li Jinfang also explored the Kam-Tai substratum in Cantonese. He stated, “The most obvious evidence of language contact is the historical substratum of the ‘defeated’ language left in the ‘victory’ language. Then, did Kam-Tai languages also leave a historical substratum in the earliest dialects of Chinese spread to the south, i.e. Cantonese and the Pinghua vernacular? The answer is definitely affirmative, as demonstrated by other researchers.” (Li 1990: 123–136)

2.2.6 In *The Manchurian Substratum and the Origins of the Neutral Tone and the Retroflex Suffixation in Beijing Mandarin* (1996), Zhao Jie researched the Manchurian substratum in modern Beijing speech by exploring word-formation and phonological patterns blending Manchurian and Mandarin in the origins of the neutral tone and retroflex suffixation. Moreover, the book also devotes a chapter to analyzing and discussing “Cases of the Manchurian Substratum in Beijing Mandarin and Patterns Integrating Manchurian and Mandarin Words”. (Zhao 1996: 1–10)

2.2.7 In “A Study of the Substratum in Southeastern Dialects of Chinese” (2005), Li Rulong

stated that there exists a Kam-Tai substratum in the southeastern dialects of Chinese. He said, “Words unique to southeastern dialects can find corresponding words of similar pronunciation and meaning in today’s Kam-Tai languages (descendants of Old Bai-Yue). In the southeastern dialects, the glottal stop, the /t-/ and /th-/ pronunciation of the Jing (精) and Qing (清) initials, and the /h-/ pronunciation of the Tou (透) initial are remnants showing different degrees of influence from Kam-Tai languages. These substrate phenomena should be formed from the Old to the Middle period.” Li also suggested: “Southeastern dialects of Chinese have been in contact with Kam-Tai languages for over two thousand years, and some substrate words may have entered into the most common variety of Chinese as used in ancient times.” (Li 2005: 1–15)

2.3 By reviewing individual studies of the last three decades, we can see that the substrate theory has achieved much progress in the study of China’s ethnic languages. The substrate theory was introduced to Chinese scholars, who subsequently found out that there are substratum languages in China. Chinese researchers started to think about the issue of substrata in Chinese languages and broadened their horizons of research. However, we also see that for most languages which have been suggested to have substrata there is insufficient scientific proof. Up to now, we have not yet found a publication that has discussed the substrate theory thoroughly and systematically. When speaking of the nature of a language with substratum elements, researchers often use uncertain words like “maybe”, or just mention the more general term “mixed language” instead of more specifically speaking of a “substratum”. As for the non-native elements in a language, some scholars simply call them a substratum without giving any reasons. In addition, there is a tendency to make an even wider application of the notion of the language with substratum elements. Consequently, substratum theory has not made even more impressive development, lingering for a long time in the stage of assumption. This situation is caused by both the complexity of the languages and insufficient research into the substrate theory.

3. Difficulties and Challenges that the Substrate Theory Encounters

3.1 Chinese scholars of ethnic languages are faced with some problematic issues in the application of the substrate theory. These issues include: (1) how to ascertain what a language with substratum elements is, (2) how to distinguish a language with substratum elements from a mixed language, and (3) how to differentiate substrate elements from elements of the adopted language.

3.2 How to Ascertain what a Language with Substratum Elements is

How to ascertain whether a language has substratum elements? The answer to this question is foundational to the study of languages with substratum elements. What is a language with substratum elements? According to the definition that I have given above, a language with substratum elements is the adopted language containing elements of the previous language. Therefore, since there is no language adoption, the Hezhou variety of Chinese cannot be considered as such. Based on published data, the Hezhou vernacular still belongs to the Chinese

language family, even though it shares some common or similar characteristics with Altaic languages, such as the case system. However, there is no evidence whatsoever to prove that people speaking Hezhou today used to speak Altaic before switching to Chinese. So, Hezhou speech can only be regarded as a regional variety of Chinese that has been affected by Altaic; it is not even a mixed language. If Hezhou speech was to be studied as a language with substratum elements, elements of language contact would be mistaken for substrate elements.

Whether a language has substratum elements or not depends on the genetic affiliation of the language. The Huihui vernacular spoken by Hui (or Muslim) people in Hainan Province has a number of important characteristics common with the Rade dialect of the Cham branch of the Indonesian group. The characteristics include: a large number of cognates, obvious correspondence in pronunciation, and major grammatical attributes. However, the Huihui vernacular also shares with Kam-Tai some common characteristics, such as “words of relation”, and some phonological and grammatical features. Zheng Yiqing, an expert on the Huihui variety, concluded: “There is no doubt that the Huihui vernacular has genetic relationship with several languages of the Cham branch including Rade. ... From the perspective of typology, it is undeniable that Huihui vernacular is similar to Dong-Dai languages. ... Therefore, we believe that the Huihui variety is a special language belonging to the Cham branch of the Indonesian group of the Austronesian family.” (Zheng 1997: 133) After systematically comparing the Huihui vernacular, Rade and Dong-Dai languages, Zheng Yiqing confirmed that Huihui speech still keeps the main features of the Cham branch of the Indonesian group of the Austronesian family. It is not language shift, but a special language belonging to the Cham branch. I believe that that conclusion is acceptable.

3.3 How to Distinguish a Language with Substratum Elements from a Mixed Language

The so-called “mixed language” is an unclear concept. Not having a consistent understanding of the term, linguists do not agree as to whether mixed languages even exist. If mixed languages do exist, then a mixed language is a new language comprising elements of two or more languages. But as to how “mixed” a language has to be before it is considered a mixed language, scholars do not have a consensus. There should be two types of mixed languages. The first type is due to language contact. In other words, the recipient language borrows words from the source language, thereby mixing elements of the recipient language and the source language. As for the proportion of elements of the two languages, there has not been any agreement. The second type is due to language shift, that is when elements of the adopted and elements of the previous language mix together, forming a new language. This kind of mixed language is a language with substratum elements. Therefore, not all mixed languages contain substratum elements; only the second type do.

Many scholars believe that Cantonese of Guangdong is a language with substratum elements. When ancient Zhuang-Dong people stopped using their original Kam-Tai languages and adopted Chinese, elements of Kam-Tai languages were brought into Chinese. This is a very plausible hypothesis because there is linguistic evidence and historical proof of the distribution of ethnic groups and migration routes. However, it is not easy at all to prove that Cantonese-speaking people had gone through a language shift in ancient times.

Do Kam-Tai languages constitute substrata of Southern Chinese dialects? To answer

this question, the first step is to solve the problem of the affiliation of Kam-Tai languages. There is a great divide concerning this in the scholarly community. One viewpoint is that the Kam-Tai languages are genetically related to Chinese. If this viewpoint can be established, Kam-Tai languages cannot be seen as substratum languages. They are at best languages greatly influenced by Chinese. The other viewpoint is that there is no genetic relationship between Kam-Tai and Chinese. It is language contact later on that causes the similarities between them. If this viewpoint is confirmed, then there are two possible explanations for the similarities. First, continually influenced by Chinese, Kam-Tai languages have changed fundamentally even in their basic structure, although still retaining some of their original characteristics. Second, Zhuang-Dong people shifted to Chinese, bringing in elements of the original languages, thereby creating a new language. This new language is a language with substratum elements.

It is not easy at all to prove that Southeastern Chinese dialects constitute an example of languages with substratum influence because it is difficult to prove there once was a language shift from Old Yue to Chinese.

3.4 How to Distinguish Substrate Elements from Elements of the Adopted Language

Coexisting in a single linguistic system, the substrate elements and elements from the adopted language would change a lot and become beyond recognition in the long historical process of evolution. It is even more difficult if the two languages are genetically related since there are shared elements descended from the proto-language. It is hard to determine which elements were brought from the previous language and which elements were from the adopted language. The further back in history a substratum was formed or when language shift took place, the more difficult it is to ascertain the exact nature of the elements involved.

4. How to Strengthen the Establishment of the Substrate Theory

The cases above show us that substrate theory is a useful theory in accounting for relations between languages, but it is just not easy to use it well. The main reason is that a substratum may have been formed very early in history and kept on changing afterwards alongside the evolution of the adopted linguistic system, making it rather difficult for scholars to recognize the substrate elements. It is essential to strengthen the theory and methodology on the substratum.

4.1 To Conduct Rigorous Comparison Starting from the Bottom

To study a language with substratum elements is to ascertain the existence of substrate elements, the relationship between substrate and non-substrate elements, and the historical development of substrate elements. To solve these problems, linguistic comparison should be carried out rigorously starting from the bottom of the genetic family tree. The true nature and characteristics of the substrate elements can only be determined based on the results of this comparison. Without the results, the genuine nature of substrate elements would not be recognized. For example, the reason why we can be sure that the Huihui vernacular is not a language with substrate elements is that researchers have conducted in-depth comparison of

Huihui speech, the Indonesian branch and Kam-Tai languages to clarify the synchronic relations among these languages.

4.2 To Examine Substrate Elements within the Linguistic System and to Avoid Using a Single Piece of Evidence

As a linguistic system, the different structural components and elements of a language are linked together as a whole. Therefore, the nature and source of the elements can be ascertained based on the mutual constraining relations of the entire system. To recognize substrate elements, we cannot just rely on the sameness or similarities of an individual element, because it constitutes only a single piece of evidence. Of course, we cannot deny an individual clue, especially since it can lead to other clues. The main point is that we cannot just stay on the level of individual clues, and certainly not base our judgment on a single point. In the past, every now and then we would find that the relationship of substrate, loan and cognate elements were determined too rashly, naturally resulting in wrong conclusions.

4.3 To Balance Evidence from Linguistic Data and from Social and Cultural History

The issue of substrata falls within the fields of both historical linguistics and sociolinguistics. The former regards a language with substratum as a product of history, to be studied by the diachronic comparative method, whereas the latter believes that the issue of substrata is constrained by social conditions and factors, so it is necessary to consider social features (including ethnicity, history, demography, geography and culture). However, it should also be emphasized that linguistic evidence is primary, while social and cultural proofs are secondary. It is inappropriate that some previous studies arrived at their conclusions based only on social and cultural relations.

References

- Chen, Wei and Chen, Peng (trans.)
 1969 *Chinese Translation of Papers on Language Substratum* (internal publication for reference within PRC).
- Gao, Mingkai
 1995 *On Languages*. Beijing: Kexue Chubanshe.
- Language Teaching and Research Office, Peking University
 1960 *Linguistics Terminology*. Beijing: Commercial Press.
- Li, Jinfang
 1990 On the voicing of voiceless fricative initials in languages in Southern China. *Studies on the Relations between Chinese and Minority Languages*, edited by the Editorial Section of the Journal of Central University for Nationalities, Beijing.
- Li, Rulong
 2005 A study of the substratum in the southeastern dialects of Chinese. *Minzu Yuwen* 5: 1–15.
- Luo, Meizhen
 1983 A Tentative View Concerning the Genetic Classification of the Kam-Thai Languages. *Minzu Yuwen* 2: 30–40.

Ma, Shujun

- 1984 Some Comparisons between the Hezhou Dialect of Chinese and Altaic Languages. *Minzu Yuwen* 2: 50–55.

Ouyang, Jueya

- 1989 Elements of Old Yue in Cantonese. *Collection of Papers on Languages and Scripts: In Celebration of Mr. Wang Li's 50 Years of Academic Study*, Beijing: Zhishi Chubanshe.
1991 Study of national minority languages with the substratum theory. *Minzu Yuwen* 6: 23–29.

Zhao, Jie

- 1996 *The Manchurian Substratum and the Origins of the Neutral Tone and the Retroflex Suffixation in Beijing Mandarin*. Beijing: Yanshan Chubanshe.

Zhao, Xiangru and Haxim

- 1981 On the Numerals of the Aynu Language and Its Essence, with Remarks on the Nature of Aynu. *Minzu Yuwen* 2: 44–48.

Zheng, Yiqing

- 1997 *The Huihui Vernacular*. Shanghai: Far East Press.